

センサ・カメラ併用型屋内測位・姿勢推定システムへの CNN 応用の提案

立川 剛至[†] 新井 イスマイル[†]

[†] 明石工業高等専門学校 電気情報工学科

1 はじめに

スマートフォンと同様の多種センサを搭載したメガネ型端末の到来により、実世界の映像に付加的な情報を重畳表示する拡張現実感 (Augmented Reality: AR) を用いたシステムに注目が集まっている。中でも、位置情報を用いた AR アプリケーションは、Ingress¹等のゲームや Skully AR-1²等のナビゲーションを必要とする分野で活発に開発されており、商業的に高い価値を持っていることが分かる。AR を実現する手法の一つとして、メガネ型端末上のスクリーン座標系で現実世界と仮想物体を重ね合わせる方法がある。しかし、全世界座標系の上で仮想物体を表示させる場合には、高精度な自己位置・姿勢推定は必要不可欠な技術となる。

屋外環境では、GPS(Global Positioning System) によって高精度な測位が可能だが、屋内環境では GPS が利用できない。そこで、RFID(Radio Frequency Identifier)、iBeacon といった機器を利用した研究が行われているが、機器の設置コストが大きいという問題がある。また、歩行者デッドレコニング (PDR) のようなジャイロセンサ・加速度センサ・地磁気センサを利用した研究も行われているが、センサ性能および保持姿勢の変動の問題がある。以上より、屋内における自己位置・姿勢推定手法は未だ確立されていない。

一方で、画像とその他のセンサを組み合わせた位置・姿勢推定手法が提案されているが、画像による位置推定の性能や画像の特徴点の量による処理速度の性能が課題となっている。そこで、CNN(Convolutional Neural Network) を用いることで、高精度な画像分類による位置推定の精度向上や、特徴点の量に影響されないことによる処理速度向上が期待できるシステムを提案する。

2 関連研究

本研究のベースとなる、Wi-Fi フィンガープリンティングとマルチセンサ、撮影画像を組み合わせた屋内測位の研究 [1] を挙げる。前準備として、対象空間の任意の位置で得られる Wi-Fi 基地局の信号強度 (RSSI) のデータベースと、画像のデータベースを構築する。位置推定を行う際は、現在地の RSSI 値を測定し加速度センサによって被測定者の一步を検出し、一步毎に歩幅推定、姿

勢推定、RSSI の測定を行う。また必要に応じて画像を撮影する。撮影画像の SIFT 特徴量を画像データベースから近傍探索することで位置推定を行う。これらの測定値からパーティクルに重みを付け重心を算出して位置と姿勢を推定する。継続撮影の場合、位置推定は平均誤差 2.5 [m] 以内に収めることに成功している。しかし、画像と PDR を組み合わせた位置推定の平均誤差は Wi-Fi と PDR を組み合わせた位置推定の平均誤差を上回っており、Wi-Fi と画像、PDR を組み合わせた位置推定結果は Wi-Fi による位置推定に大きく依存していると考えられる。よって、画像処理の改善を行うことで位置推定の精度の向上が期待できる。

撮影画像と画像マッチング手法を工夫して高精度・高速な位置推定の実現を目指している研究として、全方位カメラを用いた屋内測位の研究 [2] がある。全方位カメラを用いることで、撮影回数を減らすことができるため設備負担を小さくすることができ、マッチング範囲が広がるためマッチング点増加による位置推定の精度向上が期待できる。また、画像間の特徴点の探索に最近傍探索法である局所性鋭敏型ハッシュ (LSH) を用いることで高速な探索を実現している。しかし、この研究ではマッチングの対応点の数が場所によって大きく変動していること、マッチングが失敗しやすい場所が存在することから、適切な特徴量抽出を行えているとは言い難い。

最後に、画像認識の研究として CNN を用いた画像認識の研究 [3] を挙げる。ImageNet LSVRC と呼ばれる画像認識のコンテストにおいて、既存の特徴量抽出アルゴリズム (SIFT, SURF 等、画像の輝度値の勾配を利用) を用いた手法よりもエラー率が 10% 下回るという結果を出している。CNN は、入力に画像の画素値を取り、畳み込み層とプーリング層を繰り返し接続し、クラス分類するための学習器を最後に接続したニューラルネットワークである。そして、出力と教師データから誤差逆伝播法を用いて各層のパラメータを学習することにより、既存のアルゴリズムより良い画像の特徴量を得ることを可能としている。CNN は、コンテストにおいて良い結果を出しているが、コンテスト以外の使用事例がほとんど無く、実環境での性能が十分に検証されていない。

以上のことをまとめると、画像の特徴量抽出の性能の改善が、画像を用いた位置・姿勢推定における課題となる。

¹ <https://www.ingress.com/>

² <http://www.skully.com/>

3 提案

本研究では、我々がこれまでに地図と全天球映像を整備した大阪・梅田周辺地下街をフィールドとして検証する。地下街には、店の看板や標識、床や天井の模様と言った、人間の視覚上で自分がいる位置が認識できる特徴量を持った物体が多く集まっている。そこで、人間の脳の視覚野の処理モデルを参考にしている CNN を用いることで、画像データベースから撮影画像の分類の精度向上の期待ができる。また、CNN では特徴点を抽出して画像マッチングアルゴリズムを用いて特徴点同士を対応付けるという手順が必要無い。学習したニューラルネットワークに撮影画像を入力して設定された各層毎の上で各ピクセル毎に演算を行うため、マッチングの対応点に依存せずに位置・姿勢推定を行うことができる。よって、特徴点の量に影響されないことによる処理速度の高速化が期待できる。

図 1 に、本研究で提案する手法の流れ図を示す。データベース事前準備では、データベース構築と CNN の学習を行う必要がある。ある地点における全天球画像から取り出した任意の視点の画像の解像度を 256×256 にダウンサンプリングしたものをテンプレート画像とする。データベースには、測定位置 (世界座標系)、測定位置での観測 AP (BSSID) と RSSI のリスト、テンプレート画像 (画像情報、特徴量、方位角・仰角、奥行き情報)、ある人物に対しての歩幅を設定するためのモデル (Step モデル) を格納する。図 1 の処理の流れを以下に説明する。

1. 加速度センサによって一步を検出する。
2. 次の処理を行う。
 - (a) RSSI を測定し、BSSID と RSSI のリストを格納しているデータベースから k-NN 法を用いて位置推定を行う。
 - (b) 加速度センサ・地磁気センサ・ジャイロセンサから方位角の推定を行う。
 - (c) Step モデルと加速度センサから歩幅の推定を行う。
 - (d) メガネ型端末のカメラの撮影画像を CNN に入力し、位置、方位角、仰角の推定を行う。
3. 前回の step 検出地点での、歩幅と方位角、位置の推定値から粒子に重みを加え、重心から現在地点の位置、方位角を求める。

CNN の学習に用いるデータセットは、一つの地点 (一枚の全天球画像) につき方位・仰角・スケールを考慮して、 $360 \times 180 \times 5 = 324000$ 枚の画像を用意する。また、位置を推定する学習器、方位角を推定する学習器、仰角を推定する学習器の 3 つの CNN をそれぞれ構築し、

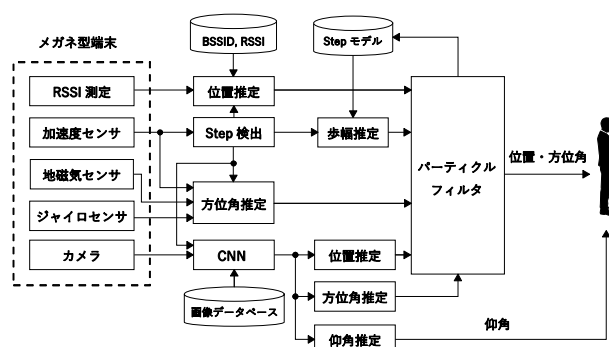


図 1 提案手法の流れ図

教師あり学習を行う。データベースに格納した各値にラベルを割り振り、各ラベルに対して正しい場合を 1、正しくない場合には 0 を割り当て、多クラス分類を行う。CNN の実装に関しては、Caffe³等のライブラリを用いる予定である。

4 おわりに

画像における特徴量は CNN によって学習できること、画像を用いた位置・姿勢推定に CNN を用いることで、画像分類の精度向上、処理速度の性能の向上が期待できることが分かった。そこで、本研究では CNN を特徴量抽出とした屋内測位の手法を提案した。今後は、CNN の各層のパラメータの調整についての調査を行う。また、実際に現地での測定を行い、手法の検証を行う。

謝辞

本研究の一部は、JSPS 科研費 25730066 の助成によるものである。

参考文献

- [1] Plamen Levchev, Michael N. Krishnan, Chaoran Yu, Joseph Menke, and Avidesh Zakhori. Simultaneous Fingerprinting and Mapping for Multimodal Image and WiFi Indoor Positioning. In *2014 International Conference on Indoor Positioning and Indoor Navigation*, pp. 6B-1, Busan, Korea, October 2014.
- [2] 畑田晃希, 河治寿都, 山崎俊彦, 相澤清晴. 全方位カメラによる位置参照画像群を用いた屋内位置推定: デジタルミュージアムでの鑑賞者の行動記録に向けて. 信学技報. MVE, マルチメディア・仮想環境基礎, Vol. 109, No. 466, pp. 57-61, March 2010.
- [3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25*, pp. 1097-1105, 2012.

³ <http://caffe.berkeleyvision.org/>